人口問題研究 (J. of Population Problems) 52−2 (1996.7) pp.17〜29

A Kinship Model Based on Branching Process

Toru SUZUKI

Introduction

Pullum (1982) developed a kinship model by which one can derive detailed frequency distribution of various kin categories. Since the model is developed on the basic branching process (Harris, 1963), it does not have the age structure and produces only eventual numbers. On the other hand, his model can provide not only the mean but also the higher moments for any distant kin in a relatively simple way.

The first section in this article is devoted to reviewing the essence of Pullum's model. Pullum showed how frequencies of female kin are calculated by using branching process.

The second section examines extension of the model to two sexes. There is a misleading assertion in Pullum's article. I will show that the relation of two sex variance to one sex variance is not so simple as suggested by him.

In the third and fourth sections, I will attempt some extensions of Pullum's model. The assumptions on marriage implied in the model are, i) the number of siblings is independent between spouses, ii) there is no remarriage. I will try to relax these assumptions in limited ways.

For the latter issue, a general model of remarriage has been developed by Goldstein (1994). I will limit myself to specific forms of remarriage and obtain such basic moments as the mean and variance.

1. One Sex Model

The starting point of the female one sex branching process is the probability distribution in the eventual number of daughters.

$$f_k = Pr\,(k\ \ daughters) \qquad (1)$$

The generating function is useful to obtain various moments of the number of daughters.

$$f(s) = \sum_{k=0}^{\infty} f_k s^k \qquad (2)$$

The mean number of daughters, or net reproduction rate (NRR), is given as the differential of the generating function with $s=1$.

$$N = f'(1) \qquad (3)$$

The second derivative of the generating function with $s=1$ gives the expectation of k^2-k. Thus, the variance of daughters is,

$$\sigma_N^2=f''(1)-[f'(1)]^2+f'(1). \tag{4}$$

To obtain the distribution of lateral kin, we need to prepare the same set for sisters. The number of sisters can be expressed with the number of daughters seen from the mother. For ego to have k sisters, her mother needs to have $k+1$ daughters. If ego is randomly chosen from generation i, which size is denoted by Z_i, Z_{i-1} gives the number of all potential mothers. Since the number of egos produced by mothers who eventually had $k+1$ daughters is $Z_{i-1}f_{k+1}$, and the ratio of daughters to mothers is the mean number of daughters, probability of k sisters is given as follows :

$$g_k=\frac{Z_{i-1}}{Z_i}f_{k+1}(k+1)=\frac{1}{N}f_{k+1}(k+1) \tag{5}$$

Substituting this into the definition of the generating function, sisters' generating function can be expressed as follows.

$$g(s)=\sum_{k=0}^{\infty}g_k s^k=\frac{1}{N}f'(s) \tag{6}$$

Because of this relationship, a moment of sisters has the one degree higher moment of daughters. For example, the mean of sisters S contains the variance of daughters.

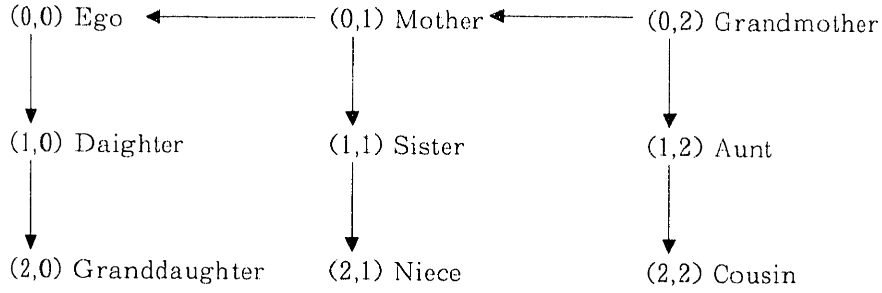
$$S=\frac{\sigma_N^2}{N}+N-1 \tag{7}$$

The variance of sisters contains the third moment of daughters.

$$\sigma_S^2=\frac{\kappa_N^3}{N}+\frac{\sigma_N^2}{N}\left(N-\frac{\sigma_N^2}{N}\right) \tag{8}$$

Pullum showed that the mean and variance of every kin category can be expressed with N, σ_N^2, S, and σ_S^2 because generating function for each category can be obtained easily by nestings of $f(s)$ and $g(s)$. It is convenient to work on Atkins' scheme for various kin categories (Figure 1). Each arrow is drawn from mother to daughter. Each vertex is labelled with i and j, where j is the generational difference between the common ancestor and ego, and i is that between the common ancestor and the kin in concern.

Figure 1 Atkins' (i, j) Lattice for Female Kin



For direct ancestors ($j=0$), the eventual number is always one since this is a one sex model. The generating function for these categories is a constant, $f(s)=s$. For direct descendants ($j=0$), a primary theorem of branching process (Harris, 1963, p.5) tells that generating functions are nestings of $f(s)$. Thus, the generating function for granddaughters is $f[f(s)]$, that for great-granddaughters is $f[f[f(s)]]$, and so forth.

Lateral kin with the same i shares the same generating function. If a lateral kin is a daughter of one of direct ancestors, then $i=1$ and generating function is $g(s)$. Granddaughters of direct ancestors have generating function $g[f(s)]$, great-granddaughters have $g[f[f(s)]]$, and so forth.

In this way, one can get generating function for any distant kin. This means that one can get any moments for all kin categories. Table 1 is the summary of mean and variance expressed with N, σ_N^2, S, and σ_S^2.

Table 1 Moments in One Sex Model

Kin	Generating Function	Mean	Variance
Mother	s	1	0
Grandmotehr	s	1	0
Daughter	$f(s)$	N	σ_N^2
Granddaughter	$f[f(s)]$	N^2	$N(N+1)\sigma_N^2$
Sister	$g(s)$	S	σ_S^2
Aunt	$g(s)$	S	σ_S^2
Niece	$g[f(s)]$	NS	$N^2\sigma_S^2+S\sigma_N^2$
Cousin (female)	$g[f(s)]$	NS	$N^2\sigma_S^2+S\sigma_N^2$

2. Extension to Two Sexes

It is easy to generalize the model to two sexes by using the fact that the sex composition of given number of children distributes binomially[1]. Let γ be the probability that a child is female. Given the number of children of both sexes, denoted as C, the number of daughters D has the binomial distribution with C and γ as parameters.

$$Pr(D=d|C=c)=\binom{c}{d}\gamma^d(1-\gamma)^{c-d},\quad D|C\sim Binomial(C,\gamma) \tag{9}$$

Then, the product γC gives the conditional mean $E(D|C)$. The unconditional mean of daughters $E(D)$ is the mean of this conditional mean. It turns out that there is a simple relationship between one sex and two sex means.

$$E(C)=\frac{1}{\gamma}E(D)=\frac{1}{\gamma}N \tag{10}$$

It is nice that this simple relation holds for any kin. Two sex mean is always some multiplier times one sex mean, as shown in Table 2. And as Pullum stated (1982, p.555), the multiplier can be obtained easily from i and j in Atkins' scheme.

$$k_{ij}=\begin{cases}2^j & \text{for } i=0 \text{ (direct ancestors)}\\ (1/\gamma)^i & \text{for } j=0 \text{ (direct descendants)}\\ 2^{j-1}(1/\gamma)^j & \text{otherwise (lateral kin)}\end{cases}$$

1 The assumption is that there is no parental control on the sex of a newborn, which guarantees the sex of each child to be identical and independent Bernoulli event.

Pullum went further and wrote as follows (1982, p.556).

"For moments of order m about the mean, $m>1$, the multipliers will simply be

$$k_{ij}=\begin{cases} 0 & \text{for } i=0 \\ (1/\gamma)^{mi} & \text{for } j=0 \\ 2^{mj-m}(1/\gamma)^{mj} & \text{otherwise.''} \end{cases}$$

This gives an impression that two sex moment is always as simple as the multiplier times the one sex moment. Let us check this for variance. Since the variance is the second order moment about the mean, m equals to 2. If the simple relation holds, the two sex variance of children should be equal to $(1/\gamma)^2$ times the one sex variance because $i=0$ and $j=0$ for children.

Recall that the variance in binomial distribution is given by $\gamma(1-\gamma)C$. This is the conditional variance of daughters given the number of children of both sexes, and can be written as $var(D|C)$. Recall that the formula which relates conditional and unconditional variances is,

$$var(D)=var[E(D|C)]+E[var(D|C)].$$

By applying this, the relationship between one sex and two sex variances turns out to be as follows.

$$var(C)=\frac{1}{\gamma^2}[var(D)-(1-\gamma)E(D)]=\frac{1}{\gamma^2}[\sigma_N^2-(1-\gamma)N] \quad (11)$$

Thus, the relation is not as simple as suggested by Pullum even for such direct kin as children. For more distant kin, the relationship becomes much more complicated as shown in Table 2. Two sex moments are always expressed with asterisks.

Table 2 Moments in Two Sex Model

Kin	Mean	Variance
Parents	2	0
Grandparents	4	0
Child	$N^*=\frac{1}{\gamma}N$	$\sigma_N^{*2}=\frac{1}{\gamma^2}\{\sigma_N^2-(1-\gamma)N\}$
Grandchild	$N^{*2}=\frac{1}{\gamma^2}N^2$	$N^*(N^*+1)\sigma_N^{*2}=\frac{1}{\gamma^4}N(N+\gamma)\{\sigma_N^2-(1-\gamma)N\}$
Sibling	$S^*=\frac{1}{\gamma}S$	$\sigma_S^{*2}=\frac{1}{\gamma^2}\{\sigma_S^2-(1-\gamma)S\}$
Uncle/Aunt	$2S^*=\frac{2}{\gamma}S$	$2\sigma_S^{*2}=\frac{2}{\gamma^2}\{\sigma_S^2-(1-\gamma)S\}$
Nephew/Niece	$N^*S^*=\frac{NS}{\gamma^2}$	$N^{*2}\sigma_S^{*2}+S^*\sigma_N^{*2}=\frac{N^2(\sigma_S^2-S)+\gamma S(\sigma_N^2+N^2)-\gamma(1-\gamma)NS}{\gamma^4}$
Cousin	$2N^*S^*=\frac{2NS}{\gamma^2}$	$2(N^{*2}\sigma_S^{*2}+S^*\sigma_N^{*2})=2\frac{N^2(\sigma_S^2-S)+\gamma S(\sigma_N^2+N^2)-\gamma(1-\gamma)NS}{\gamma^4}$

This fact does not change the conclusion on the correlation between frequency of kin in Pullum and Wolf (1991). They showed that, in homogeneous and independent case as in the stable population, correlation exists only between direct descendants, and the correlation coefficient is simply a function of the mean number of daughters. For example, the correlation between the frequency of daughters and granddaughters is expressed as follows (p.397).

$$r=\sqrt{\frac{N}{N+1}} \tag{12}$$

Because only the mean matters, their conclusion is not affected by the relationship between one sex and two sex variances.

3. Dependence in the Number of Siblings

In the two sex version of Pullum's model, the independence between the sibling frequency of husband and that of wife was assumed. This means that a woman marries a man regardless of how many siblings he has, and vice versa. Two ideal cases in which this condition is dropped is discussed here to see how marriage pattern can alter the distribution of lateral kin.

First, assume that a husband always has the same number of siblings as his wife has. In this case of perfect homogamy, the number of uncles and aunts of both lines is simply twice as much as that of maternal line. If K is the number of siblings of a mother, $2K$ is the number of all uncles and aunts seen from her child. The basic theorems on the multiplication of a random variable show that the mean is $2S^*$ and the variance is $4\sigma_N^{*2}$.

Let J be the number of cousins. Since the model assumes the homogeneous and independent reproductive behavior, the conditional mean and variance of J given K are $2KN^*$ and $2K\sigma_N^{*2}$. Using the basic theorems on conditional moments, we can show that the unconditional mean of J is $2N^*S^*$ and the unconditional variance is $4N\sigma_N^{*2}+2S^*\sigma_N^{*2}$. Comparing these moments with those in Table 2, it can be shown that the perfect homogamy makes no change with mean but raises the variance of lateral kin.

Second, let us see what happens if a woman with no siblings never marries with a man with no siblings. A marriage between sole children may cause problems, especially in traditional settings. In such a case, one of family names will die out. In addition, the young couple may not be able to take care of parents of both spouses, because there is no sibling to share the task of supporting elderly.

If marriages between sole children are strictly excluded, then every member of the society has at least one uncle or aunt ever born. This can be seen as a simple problem of conditioning.

Let g_k^* be the probability function of sibling frequency, the two sex version of g_k. Let I be an indicator variable which takes one when the marriage is "valid", namely, at least one of couple has at least one sibling. Then, the probability of $I=0$ is that of sibling frequency is zero for both spouses, which is g_0^{*2}.

Let X be the frequency of uncles and aunts. As in Table 2, the mean number of uncles and aunts is $2S^*$ in the case of independence. This unconditional mean should be the mean of conditional means conditioned by I.

$$2S^{\cdot}=g_0^{\cdot 2}E\,(X|I=0)+(1-g_0^{\cdot 2})\,E\,(X|I=1)$$

Since $E\,(X|I=0)$ is the mean number of uncles and aunts when neither parent has any siblings, it is zero. Another conditional mean $E\,(X|I=1)$ is the goal here, the mean of uncles and aunts when marriages between sole children are excluded.

$$E\,(X|I=1)=\frac{2S^{\cdot}}{1-g_0^{\cdot 2}} \tag{13}$$

To get the conditional variance $var\,(X|I=1)$, it is necessary to know the variance of conditional means.

$$var\,[E\,(X|I)]=g_0^{\cdot 2}\,(0-2S^{\cdot})^2+(1-g_0^{\cdot 2})\,(\frac{2S^{\cdot}}{1-g_0^{\cdot 2}}-2S^{\cdot})^2=\frac{4S^{\cdot 2}}{1-g_0^{\cdot 2}}-4S^{\cdot 2}$$

Using the fact that $var\,(X)=2\sigma_S^{\cdot 2}$ and $var\,(X|I=0)\;0$, and applying the formula $var\,(X)=var\,[E\,(X|I)]+E\,[var\,(X|I)]$,

$$var\,(X|I=1)=\frac{2\sigma_S^{\cdot 2}+4S^{\cdot 2}}{1-g_0^{\cdot 2}}-\frac{4S^{\cdot 4}}{(1-g_0^{\cdot 2})^2}. \tag{14}$$

This is the variance of uncles and aunts when marriages between sole children are excluded. The same conditioning can be done for the frequency of cousins. The mean and variance of cousins after the exclusion of invalid marriages would be as follows :

$$E\,(X|I=1)=\frac{2N^{\cdot}S^{\cdot}}{1-g_0^{\cdot 2}} \tag{15}$$

$$var\,(X|I=1)=\frac{2\sigma_S^{\cdot 2}N^{\cdot 2}+2\sigma_N^{\cdot 2}S^{\cdot}+4N^{\cdot}S^{\cdot}}{1-g_0^{\cdot 2}}-\frac{4N^{\cdot 2}S^{\cdot 2}}{(1-g_0^{\cdot 2})^2} \tag{16}$$

It is easy to see that conditional means are always greater than the original means. For the conditioning to be meaningful, $g_0^{\cdot}$ should be greater than zero. Then, $1/\,(1-g_0^{\cdot 2})$ in (13) or (15) works as an inflater. On the other hand, conditional variances can be greater or smaller than the original variances. It depends on the shape of distribution.

4. Remarriage and Half-Siblings

An analytical model of remarriage is surprisingly difficult. One needs to make many assumptions to save desirable simplicity. I repeat four of five assumptions made by Goldstein (1994, p.7).

(1) All the births are given during a marriage of a woman.
(2) The father of a child is always the marital partner of a woman.
(3) A remarriage does not change the complete fertility of a woman.
(4) The fertility rate is the same for each marriage.

The number of marriages experienced by a woman is called "marity", by analogy with "parity", in the Goldstein model (p.7). Assumptions (3) and (4) claim the independence between marity and parity of a woman.

In addition to these assumptions, I will limit marity up to twice and exclude marriages between ever married persons to examine half-siblings from different father and that from different mother separately. Thus, instead of independence in marity assumed by Goldstein, mutual exclusiveness is assumed here.

(5) No one marries three times, namely, the maximum marity is two.
(6) The spouse of a remarried person is always first married.

Goldstein gave the probability functions of half-siblings in a general setting (1994, pp.11-12). My goal here is to derive the specific expressions for the mean and variance of half-siblings, which Goldstein did not show.

4-1. Mother's remarriage and half-siblings from different father

Let I be an indicator variable to mark a remarriage of a woman, and p_M be its probability. Then, $p_M = Pr\ (I=1)$.

If a woman married twice ($I=1$), we want to know the distribution of her children between two marriages. Let t_1 be the duration of her first marriage, and t_2 that of her second marriage. Assume that every woman bears a child according to an age independent fertility rate λ. In this case, the number of children born in the first marriage W_1 and in the second marriage W_2 are the Poisson processes.

$$\text{first marriage} \quad W_1 \sim Poisson\ (\lambda t_1)$$
$$\text{second marriage} \quad W_2 \sim Poisson\ (\lambda t_2)$$

Let θ be the relative length of the first marriage, $\theta = t_1 / (t_1 + t_2)$. It is known that the distribution of W_1 given the number of all children $W = W_1 + W_2$ becomes the binomial distribution with W and θ (Stone, 1989, p.298).

$$Pr\ (W_1 = w_1 | W = w,\ \Theta = \theta,\ I = 1) = \binom{w}{w_1} \theta^{w_1} (1-\theta)^{w-w_1} \tag{17}$$

If θ is $1/2$, namely a remarriage always equally divides her effective reproductive period, a nice property appears. A randomly chosen ego has y half siblings in either of the following two cases.

(a) $W_1 = y$ and the ego belongs to the second marriage.
(b) $W_1 = w - y$ and the ego belongs to the first marriage.

Given $W = w$, $\theta = 1/2$, and $I = 1$, summing the two probabilities above results another binomial distribution.

$$Pr\,(Y=y)=\binom{w}{y}\left(\frac{1}{2}\right)^{w}\left(1-\frac{y}{w}\right)+\binom{w}{w-y}\left(\frac{1}{2}\right)^{w}\left(1-\frac{y}{w}\right)$$

$$=\binom{w}{y}\left(\frac{1}{2}\right)^{w-1}\left(1-\frac{y}{w}\right)$$

$$=\frac{w!}{y!\,(w-y)\,!}\left(\frac{1}{2}\right)^{w-1}\frac{w-y}{w}$$

$$=\frac{(w-1)\,!}{y!\,(w-y-1)\,!}\left(\frac{1}{2}\right)^{w-1}$$

This is a binomial distribution with $W-1$ and $1/2$. Let us define $X=X-1$. Since W is the number of all the children seen from a mother, X is the number of all the siblings (whole-siblings and half-siblings) seen from an ego. The above result shows that the number of half-siblings Y has the binomial distribution with X and $1/2$.

In this case, it is relatively easy to get the mean and variance of half-siblings. The conditional mean is $E\,(Y|X,\ I=1)=X/2$. By the assumption of independence between marity and parity, averaging this conditional mean over X gives $E\,(Y|I=1)=S^{\cdot}/2$, where $S^{\cdot}$ is the mean number of siblings. By using the probability that a mother marries twice, $p_M=Pr\,(I=1)$, the mean number of half-siblings in the whole population is given as follows :

$$E\,(Y)=\frac{p_M S^{\cdot}}{2} \tag{18}$$

The conditional variance is $var\,(Y|X,\ I=1)=X/4$, because the distribution of Y given X and $I=1$ is binomial with X and $1/2$. To get the marginal variance with respect of X, it is sufficient to know that $var\,[E\,(Y|X,\ I=1)]$ is $\sigma_S^{\cdot 2}/4$ and $E\,[var\,(Y|X,\ I=1)]$ is $S^{\cdot}/4$, where $\sigma_S^{\cdot 2}$ is the variance of siblings appeared in Table 2. Thus, $var\,(Y|I=1)$ is $\sigma_S^{\cdot 2}/4+S^{\cdot}/4$. In the same manner, we can get the unconditioned variance for the whole population.

$$var\,(Y)=\frac{p_M}{4}\,[\sigma_S^{\cdot 2}+S^{\cdot}\,(1+(1-p_M)\,S^{\cdot})] \tag{19}$$

The results are such nice when $\theta=1/2$ is assumed, but it is a strong assumption that a woman switches to the second marriage at the middle of her reproductive period.

Goldstein assumed the uniform distribution on the configuration of children among marriages (1994, p.12). I will show that the Goldstein model appears when θ distributes uniformly on (0,1). The density of θ is one if uniform distribution is assumed, and the joint distribution of W_1 and θ is the same as (17). To obtain the marginal distribution W_1, by integrating (17) over θ, the following theorem on beta and gamma functions is helpful.

$$\int_0^1 \theta^m\,(1-\theta)^n\,d\theta=\frac{\Gamma\,(m+1)\,\Gamma\,(n+1)}{\Gamma\,(m+n+2)}=\frac{m!n!}{(m+n+1)\,!}$$

Applying this formula, it becomes clear that W_1 (number of children born in the first marriage) uniformly distributes on W (number of all the children).

$$Pr\,(W_1 = w_1 | W = w, I = 1) = \int_0^1 \binom{w}{w_1} \theta^{w_1} (1-\theta)^{w-w_1} d\theta$$

$$= \binom{w}{w_1} \frac{w_1!\,(w-w_1)\,!}{(w+1)\,!}$$

$$= \frac{1}{w+1}$$

This time, instead of the binomial distribution with $W-1$ and $1/2$, the probability function of half-siblings becomes as follows :

$$Pr\,(Y = y | W = w,\ I = 1) = 2\left[\frac{1}{w+1}\left(1-\frac{y}{w}\right)\right] = \frac{2\,(w-y)}{w\,(w+1)},\quad w > 0. \tag{20}$$

The formula above is actually a reduced form of Goldstein's general formula (1994, p.12) when marity is two. I will continue seeking the mean and variance of half-siblings. When both the experience of remarriage and the number of children are given, the conditional mean is,

$$E\,(Y | W,\ I = 1) = \sum_{y=0}^{W} y \frac{2\,(W-y)}{W\,(W+1)} = \frac{W-1}{3},\quad W > 0.$$

By assumption of the independence between marity and parity, the average number of half-siblings among those whose mother married twice is $E\,(Y|I=1) = S^{*}/3$. Then, the unconditional mean for the whole population is,

$$E\,(Y) = \frac{p_M S^{*}}{3} \tag{21}$$

From the equation (20), the conditional variance given the mother's remarriage and the number of her children is,

$$var\,(Y | W,\ I = 1) = \sum_{y=0}^{W} \left(y - \frac{W-1}{3}\right)^2 \frac{2\,(W-y)}{W\,(W+1)} = \frac{(W+2)\,(W-1)}{18}.$$

For those whose mother married twice, the variance is,

$$var\,(Y | I = 1) = \frac{3\sigma_{S^*}^{2} + 3S^{*} + S^{*2}}{18}$$

The unconditional variance for the total population is,

$$var\,(Y) = \frac{p_M}{18}\left[3\sigma_{S^*}^{2} + 3S^{*} + (3 - 2p_M)\,S^{*2}\right] \tag{22}$$

4-2. Father's remarriage and half-siblings from different mother

Father's remarriage causes much less problems than that of mother. Since it is assumed that both wife and ex-wife of the father are first married, the number of half-siblings is that of all children born to another wife of ego's father. Since it is also assumed that parities of different women are independently and identically distributed, the distribution of half-siblings is simply that of children.

Let I be the indicator variable of father's remarriage, $p_F = Pr\,(I=1)$ be the probability of father's remarriage, and Z be the number of half-siblings from different mother. Then, the distribution of Z given that the father married twice is the distribution of children.

$$Pr\,(Z=z\,|\,I=1)=f_z^{*} \tag{23}$$

This can also be seen as a simplification of the Goldstein model. When the marity of wife and ex-wife is fixed to m, his formula for half-siblings through father is as follows (1994, p.11) :

$$v_z=\sum_n p_z f_n^{*}$$

Here, p_Z is the probability that z of her n children are born in the marriage with ego's father.

$$p_z=\binom{n}{z}\frac{1}{m^n}\,(m-1)^{n-z}$$

However, under my simplifying assumption that $m-1$ for both wife and ex-wife, $p_z=1$ if $z=n$ and 0 otherwise. Then, $v_z=f_n^{*}$ if $z=n$ and 0 otherwise, which is equivalent with the equation (23).

Thus, conditional moments of half-siblings given that ego's father married twice are simply those of children. Namely, $E\,(Z|I=1)$ is N^{*} and $var\,(Z|I=1)$ is σ_N^{*2}. The unconditional mean for the total population is,

$$E\,(Z)=p_F N^{*}. \tag{24}$$

Knowing that $var\,[E\,(Z|I)]=p_F\,(1-p_F)\,N^{*2}$ and $E\,[var\,(Z|I)]=p_F\sigma_N^{*2}$, the unconditional variance is,

$$var\,(Z)=p_F\,[\sigma_N^{*2}+(1-p_F)\,N^{*2}]. \tag{25}$$

5. Conclusion

This article has discussed an analytical model of kinship frequencies. The frequency by kin category is important because it determines the demographic condition of household size and composition. Though household dynamics is more adequately studied through computer simulations, the availability of kin as a demographic determinant can be modeled analytically. The kin availability is also important as the source of economic and social supports. Although the long term trend shows a decline in the importance of kin in everyday life, there still exist exchanges of goods and services between parents and adult children even when they do not live together.

The kinship models also have applications to demographic method. Many kin-based measures of demographic indices have been invented. The mortality estimates with "children ever born, children dead" method formalized by Brass (United Nations, 1990) and the estimate of growth rate with the ratio of sisters by Goldman (1978) and its extension by Wachter (1980) and McDaniel and Hammel (1980) belong to this class. It will be nice, however, if the reliability of estimates can be asserted with probability theory especially when small sample data are used. In this connection, the application of branching process to the methodological issue could be interesting.

Acknowledgement

Special thanks to Joshua Goldstein. Personal communication with him revealed underlying assumptions that I was not aware of, and helped me with checking and developing my analysis.

References

Atkins, J. R., 1974
"On the fundamental consanguineal numbers and their structural basis", *American Ethnologist* 1, pp.1-31.

Goldman, Noreen, 1978
"Estimating the intrinsic rate of increase of a population from the average numbers of younger and older sisters", *Demography* Vol.15, No.4, pp.499-507.

Goldstein, Joshua R., 1994
"A demographic model of stepfamily formation, with application to 20th century American and pre-revolutionary French populations", unpublished paper.

Harris, T. E., 1963
The Theory of Branching Processes, Englewood Cliffs, NJ, Addison-Wesley.

McDaniel, C. K. and E. A. Hammel, 1984
"A kin-based measure of r and an evaluation of its effectiveness", *Demography* Vol.21, No.1, pp.41-51.

Pullum, Thomas W., 1982
"The eventual frequencies of kin in a stable population", *Demography* Vol.19, No.4, pp.549-565.

Pullum, Thomas W. and Douglas A.Wolf, 1991
"Correlations between frequencies of kin", *Demography* Vol.28, No.3, pp.391-409.

Stone, Charles J., 1989-1994
A First Course in Probability and Statistics, Volume 1 : Probability, University of California at Berkeley.

United Nations, 1990
Step-by-Step Guide to the Estimation of Child Mortality, NY : United Nations [Population Studies No.107] .

Wachter, Kenneth W., 1980
"The sisters' riddle and the importance of variance when guessing demographic rates from kin counts", *Demography* Vol.17, No.1, pp.103-114.

Abstract

A Kinship Model Based on Branching Process

Toru Suzuki

The kinship model based on branching process developed by Pullum (1982) was examined and a few minor extensions were attempted. The basic branching process can produce the distribution of direct kin in one sex model of females. Pullum introduced the generating function of sisters so that the model can give any moments for all kin categories.

Pullum also discussed the extension to two sexes with binomial distribution. He showed that the two sex mean is always simply a multiple of the corresponding one sex mean. This article, however, showed that such a simple relationship as Pullum suggested does not hold for the variance.

In the Pullum model, the independence of sibling size between spouses was assumed and relatives through a remarriage were not considered separately. For the former restriction, the moments in perfect homogamy and in exclusion of mating between persons without sibling were derived and compared with those in random marriage. It was shown that the variance of sibling size is greater in homogamy than in random marriage. If the marriage between male and female who both have no sibling is prohibited, the average sibling size reduces but the change in variance depends on the shape of distribution.

As for remarriage, Goldstein (1994) gave the generating function for half-siblings. This article showed that his model is based on the Poisson process in which the number of births by marriage depends only on the length of marriage. In addition, specific formulae for the mean and the variance were obtained by adding some stronger assumptions than those in the Goldstein model.

分岐過程にもとづく親族モデル

鈴 木 透

本稿では，Pullum（1982）による分岐過程にもとづく親族数分布モデルについて考察し，若干の拡張を試みた．単性女子のモデルで，直系親族については分岐過程の基礎理論で扱える．Pullum は娘数を与える再生産関数に加え，姉妹数を与える関数を導入し，どんな遠い親族のどんな高次のモメントをも簡単に導くことができるようにした．

Pullum は両性への拡張についても論じており，これは出生児が女児である確率 γ をパラメタとする二項分布によって行なう．この場合，平均については単性と両性の間に単純な関係が成り立つ．例えば子供，キョウダイ，オジ・オバの平均は単性の $1/\gamma$ 倍であり，孫，オイ・メイ，イトコの平均は単性の $1/\gamma^2$ 倍などとなる．しかし分散については，Pullum が示唆しているような単純な関係は成り立たないことを本稿で示した．

Pullum のモデルでは，結婚について(1)夫妻のキョウダイ数は独立，(2)再婚はない，という二つの

仮定が設けられている．本稿では限定的ながらも，これらの制限を緩和することを試みた．

夫妻のキョウダイ数については，完全同類婚の場合と一人っ子どうしが結婚しない場合について，オジ・オバおよびイトコ数の平均と分散を求め，Pullum が扱っているランダム婚の場合と比較した．その結果，完全同類婚であれば平均は変化しないが分散はランダム婚の場合より大きくなり，一人っ子どうしの結婚が禁止されると平均はランダム婚の場合より小さくなるが，分散の差異は分布に依存することを示した．

再婚による異父キョウダイ，異母キョウダイのモメントについては，Goldstein（1994）が一般的な確率母関数を与えている．本稿では，初婚・再婚それぞれにおける出生児数が結婚持続期間だけに依存するポワソン過程から Goldstein のモデルが導かれることを示した．また，Goldstein より限定的な場合について，異父・異母キョウダイの平均と分散を得る具体的な式を導出した．

分岐過程では年齢が捨象されており，出生力と死亡力を分離できないのが難点である．しかし式が単純で，確率モデルでありランダム性を扱えるという利点がある．分岐過程に限らず，親族数と人口過程を結びつける確率モデルは，さらに研究されて良い分野だろう．